The Behavior Analyst 1992, 15, 147–156 No. 2 (Fall)

Trends in the Measurement of Social Validity

Craig H. Kennedy
University of Hawaii

Since its inception in the mid-1970s, social validity has provided applied behavior analysts with a critical measure of the social impact and importance of their interventions. Recent discussion, however, has questioned the use of this construct in regard to the frequency and types of social validity measures employed in research. Despite the ensuing discussion, virtually no quantitative information has been made available to frame various perspectives and opinions. The purpose of this report is to present a content analysis of social validity measures used over the previous 20 years. Social validity was assessed along three dimensions: (a) type of assessment, (b) focus of assessment, and (c) time of assessment. Articles published in the *Journal of Applied Behavior Analysis* (1968–1990) and *Behavior Modification* (1977–1990) were surveyed. The results of the content analysis indicate that current applications of social validation procedures are presented in 20% of the articles surveyed. The majority of articles used subjective evaluation of outcomes following intervention to assess social validity. In addition, the data indicated that normative comparison was a rarely used method of social validation and that its use has been decreasing over time.

Key words: social validity, applied behavior analysis, assessment, measurement, subjective evaluation, normative comparison

---

Social validity is concerned with the social desirability and usefulness of changes in behavior. It is an attempt to go beyond "clinical judgment" to derive information from the broader social environment of the individual(s) whose behavior is being changed. As the literature pertaining to social validity has grown, it has encompassed a number of related areas, including (a) consumer satisfaction (Bornstein & Rychtarik, 1983; Lebow, 1982; McMahon, 1983), (b) treatment acceptability (Elliot, Witt, Galvin, & Moe, 1986; Gullone & King, 1989; Kazdin, 1980; Kazdin & Wilson, 1978; Singh & Katz, 1985), (c) ecological validity (Brookes & Baumeister, 1977; Gaylord-Ross, 1979; Hawkins, 1975; Scott, 1980), and (d) the clinical importance of treatment outcomes (Barlow, 1981; Lovitt, 1978; Yeaton & Sechrest, 1981).

The focus of social validity assessments rests with the question: "Is this change in behavior, and/or the process used to change it, of social value?" The original impetus for developing social validity assessments was to address concerns that the measures typically employed by applied behavior analysts were not sensitive to the broader social context of variables affected. Given the effectiveness of behavior-analytic interventions, dependent measures need to assess more than discrete antecedents, responses, and consequences. The broader social ecology within which the intervention is embedded also needs to be assessed (Rogers-Warren & Warren, 1977; Willems, 1974).

Social validity, although a recent development in applied behavior analysis, was predated by research in three areas: (a) patient satisfaction with medical treatment (e.g., Koos, 1955; Makeover, 1950; Wienerman, 1964), (b) client expectations and satisfaction with psychotherapy (e.g., Barahal, Bramer, & Shostrom, 1950; Bordin, 1955; Grant, 1954; Rogers, 1942), and (c) consumer and employee satisfaction in business (Goode & Fowler, 1949; Herzberg, Mausner, & Syndern, 1959; Roethlisberger & Dickson, 1939). Each of these disciplines developed indices that measured how individuals perceived the adequacy and/or desirability of services being provided.

Social validity appears to be evolving

The author would like to thank James Halle, Thomas Haring, and Marta Valdez-Mencheca for their comments on a previous version of this paper and Susan Fox and Tiina Itkonen for their assistance with the content analysis.

Correspondence concerning this article should be sent to Craig H. Kennedy, College of Education, 1776 University Avenue, Wist 208, University of Hawaii, Honolulu, HI 96822.

as a measure. As evidenced by a recent special issue of the *Journal of Applied Behavior Analysis* (*JABA*), the definitions and uses of social validity are in the process of expanding from the original definitions provided by Kazdin (1977) and Wolf (1978). Although the original definitions have been used in numerous investigations over the past 15 years, several authors have called for expanded assessments and/or methodological revision of social validation measures (Hawkins, 1991; Schwartz & Baer, 1991; Winett, Moore, & Anderson, 1991).

Measuring Social Validity

In its current state, research using social validity measures can be analyzed in terms of three distinct dimensions. The first focuses upon the *type* of information. The second is derived from the *focus* of information that is collected. The third dimension is based upon the *time* between intervention and the assessment process. Each dimension, and its subunits, are discussed below. The various combinations will then be used as the basis for a content analysis of research in applied behavior analysis.

Two basic strategies have been used for the collection of social validity information: (a) *subjective evaluation* and (b) *normative comparison.* Subjective evaluation is based upon individuals' (e.g., experts, relatives, teachers, students) ratings or statements regarding some aspect of the intervention (Kazdin, 1977; Wolf, 1978). Normative comparison is based upon the comparability of a person's performance before or after an intervention with a group of individuals whose behavior is considered to be typical or desirable (Van Houten, 1979). Together, these approaches have defined approaches to collecting information regarding experimental effects of the independent variable(s) that are grounded within the broader social context in which the intervention occurs.

Researchers using subjective evaluation and normative comparison have analyzed three specific aspects of the intervention process (Wolf, 1978). These are (a) selecting the goals of intervention, (b) selecting the procedures to be used in the intervention, and (c) assessing the outcomes of the intervention. Social validation can be used to *select goals* for intervention, by asking people who interact and know a person (or subject population) well, or by assessing normative rates. By doing this, it is hoped that the informants can identify critical behaviors in need of change. *Procedural validation* is a process by which the form or type of intervention is assessed (cf. Kazdin, 1980). *Outcome validation,* as its title implies, assesses the perceived changes in a participant's or other's behavior as a result of an intervention (typically in terms of perceived adequacy). As a collection, these three aspects of the behavior-change process comprise the content type of social validity assessments.

A third distinction that can be drawn regarding the use of social validation measures focuses upon their use as *pre-* versus *postintervention* measures. Any of the assessment methods mentioned above can be used prior to intervention (as selection procedures or criterion measures) or following intervention to assess the perceived/normative nature of some aspect of the behavior-change process. Each of the 12 possible combinations is described below and is subsequently employed in an analysis of the use of social validity.

Preintervention subjective evaluation of goals. This approach to social validity assessment focuses upon perceptions regarding the most appropriate behavior to be changed prior to beginning an intervention. That is, before an intervention is introduced, individuals in a person's social environment are asked what behavior(s) should be changed.

Preintervention subjective evaluation of procedures. This strategy focuses upon perceptions regarding procedures to be used in an intervention prior to the start of the analysis. For instance, Clark, Greene, Macrea, McNees, Davis, & Risley (1977) asked parents to list various strategies that they used during visits to the grocery store to limit their children's problem behavior. Clark et al. then used

the five most commonly listed approaches as the basis of an intervention package to teach parents effective techniques for dealing with their children's problem behavior while shopping.

Preintervention subjective evaluation of outcomes. The primary goal of this type of social validity assessment is to determine, using individuals' subjective statements, what a desirable target level for behavior reduction or increase should be. Essentially, this assessment tactic asks people to determine when an intervention should be considered successful. Unlike preintervention subjective evaluation of goals, which asks individuals to identify target behaviors for intervention, preintervention subjective evaluation of outcomes asks individuals to identify levels or rates at which the intervention should be considered successful.

Preintervention normative comparison of goals. This approach identifies target behaviors prior to intervention by an assessment of types of behavior among typical (or desirable) populations. Using this approach, interventionists can determine prior to treatment what desirable behavior occurs among a typical population. The determination of what behavior is prevalent and the typical form(s) it takes are often the focus of this approach to social validity assessment. By using this approach, for example, interventionists can determine what types of clothing styles and color combinations are currently "in fashion" as a means of establishing what clothing items to teach young women with disabilities to select and wear (e.g., Nutter & Reid, 1978).

Preintervention normative comparison of procedures. This approach seeks to determine intervention typologies by assessing normative levels for particular interventions prior to initiation of treatment. Green, Hardison, and Greene (1984) "turned the table on parental advice" by allowing parents to describe what practices they used to reduce the problematic behavior of children when dining in restaurants. This strategy, allowing a survey of parents to determine effective and efficient interventions, was then employed in a multiple baseline across families and restaurants with positive quantitative and subjective results.

Preintervention normative comparison of outcomes. Defining the level of target behavior prior to intervention is the focus of this approach. By prespecifying target behavior levels based upon a normative sampling strategy, typical levels of desired or undesired behavior are specified for termination of the intervention.

Postintervention subjective evaluation of goals. This approach asks, in essence, "Should we have chosen this behavior(s) as a goal for intervention?" By asking experts, typical community members, or individuals who interact with the people who are the focus of intervention if the topographies chosen for reduction are appropriate, the subjective value of the goals is assessed.

Postintervention subjective evaluation of procedures. Researchers have also found it of value to ask others for their perceptions regarding interventions used in an investigation. For example, Friman, Finney, Rapoff, and Christophersen (1985) were able to assess parents' satisfaction with a prompting strategy to keep pediatric appointments for their children. The results indicated that all 88 parents who participated in the investigation were satisfied with the prompting procedures and wanted to continue their use.

Postintervention subjective evaluation of outcomes. This strategy asks experts, parents, participants, and others the question: "Was there a perceivable change in behavior as a result of the interventions?" An example is provided by Charlop and Milstein (1989), who used videotape modeling to increase the conversational skills of children with autism across people, settings, and topics. Following intervention, parents of nondisabled children viewed videotapes of the conversations and rated the social importance of the interactions (e.g., "The child shows an interest in the conversation").

Postintervention normative comparison of goals. Determining whether the

goals selected for intervention are similar to more typical (or desirable) populations is the focus of this approach. As suggested by Van Houten (1979), this approach encompasses the types of behavior that have been changed as a result of the intervention. For instance, after teaching a young woman with severe disabilities to apply various types of make-up (e.g., lip gloss, rouge), an experimenter might observe same-age peers to assess whether the goals of instruction were actually typical of the age group (see also Nutter & Reid, 1978).

Postintervention normative comparison of procedures. A tactic similar to postintervention normative comparison of goals is the normative comparison of procedures following intervention. This approach asks whether the procedures used in an investigation are typical of the types of interventions applied by others. This strategy is similar to that employed by Green et al. (1984), except that the information is obtained after the investigation is completed.

Postintervention normative comparison of outcomes. This approach is exemplified by a study conducted by Stevenson and Fantuzzo (1986). An intervention using a self-control package to increase the arithmetic proficiency of students of low socioeconomic status was used. After the intervention had been applied and arithmetic accuracy increased, an assessment was conducted in which the typical performances of students not identified as having problems with arithmetic were compared with those of the treatment group. Thus, normative rates of behavior were obtained post hoc to assess the intervention's effects.

Next, we present a content analysis describing patterns of social validity use and discuss how these trends relate to the current debate regarding this construct. The 12 approaches derived from the dimensions of type, focus, and time of assessment will be used as a means of categorizing the use of social validity. By conducting such an analysis, this article seeks to provide researchers with a data-based description of how researchers are (or are not) using social validity measures and how the pattern of use may predict future needs and issues relating to social validity.

METHOD AND RESULTS

All research articles published in *JABA* (1968–1990) and *Behavior Modification* (1977–1990) were reviewed to assess the frequency and types of social validity measures employed. The two journals were selected because they were the original archival sources for the presentation of social validity assessments in applied research (Kazdin, 1977; Wolf, 1978). The search was conducted by reviewing all data-based research articles and scoring each along the three dimensions described previously. The articles identified through this process were included in the content analysis. Of those articles reviewed, 20% were read independently by a second person, and an overall interrater agreement for occurrences of 94% and nonoccurrences of 99% was achieved using an item-by-item approach.

Overall Use of Social Validity

A total of 125 articles published in *JABA* and 53 in *Behavior Modification* have presented social validity measures. As shown in Figure 1, only a small proportion of research reports published in *JABA* between 1968 and 1975 employed social validity measures (with the exceptions being Brigham, Graubard, & Stans, 1972; Briscoe, Hoffman, & Bailey, 1975; Fawcett & Miller, 1975; Maloney & Hopkins, 1973; O'Brien & Azrin, 1972; Phillips, Wolf, & Fixsen, 1973; Van Houten, Morrison, Jarvis, & MacDonald, 1974). Articles were based primarily upon directly observed quantities and rarely included subjective perceptions or normative sampling. This pattern changed in the mid-1970s. With an increased concern for the broader "ecological" effects of behavioral interventions (e.g., Baer, 1974; Willems, 1974) and proposals to incorporate new measures in applied research (Kazdin, 1977; Wolf, 1978), a greater proportion of articles began presenting social validity data. This trend reached a peak during the early 1980s. Following the peak in 1983, the percent-

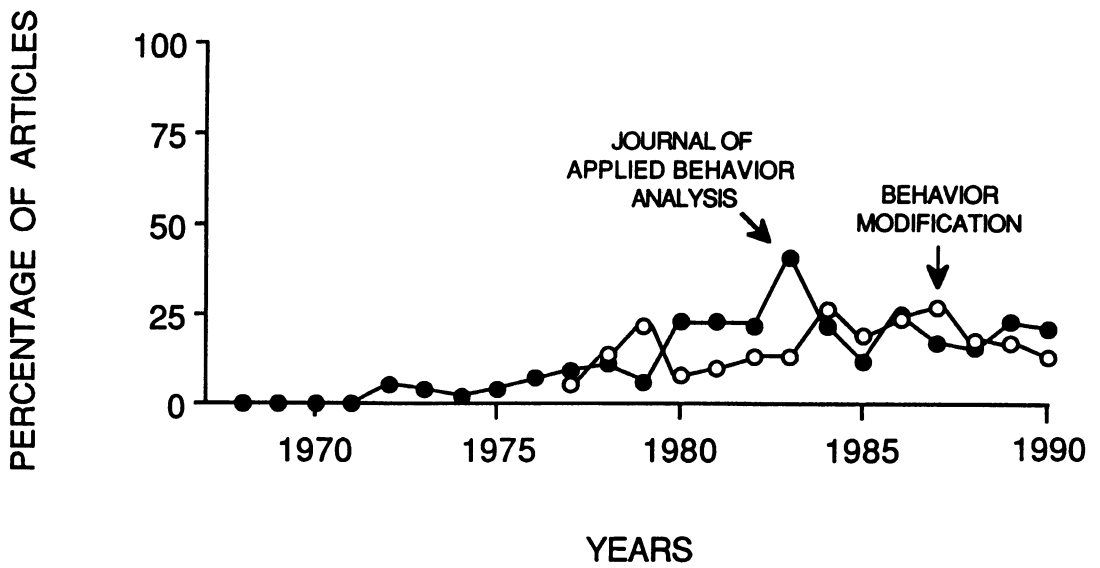

Figure 1. The use of social validity measures in the *Journal of Applied Behavior Analysis* (1968–1990) and *Behavior Modification* (1977–1990). Data are presented for the percentage of data-based articles employing social validity measures.

age of articles publishing social validity data has decreased. Currently, approximately 20% of data-based articles in *JABA* and *Behavior Modification* use social validity measures. Although this percentage has decreased since the early 1980s, trends have stabilized in recent years.

Differences in the Use of Subjective Evaluation and Normative Comparison

During the period of 1968 to 1990, the largest proportion of articles presenting social validity data used subjective evaluation (see Figure 2). For instance, in 1983, 91% of *JABA* articles presenting social validity data (41% of all articles) used subjective evaluation, whereas 9% of articles presenting social validity data (3% of all articles) used normative comparison. As Figure 2 shows, normative comparisons were rarely used, suggesting that subjective evaluation has become the almost exclusive means of assessing social validity.

Table 1 presents the frequency of pre- and postintervention assessments of goals, procedures, and outcomes. (Note that data do not pertain to specific studies, but to the number of times each type of social validity strategy has appeared in a study.) The numbers for subjective evaluation in Table 1 show that postintervention assessments were used more frequently than preintervention assessments. Of the preintervention subjective evaluations that occurred, 79% concerned the goals of intervention. Rarely were personal perceptions of procedures or outcomes assessed prior to intervention. However, 96% of postintervention subjective evaluations occurred for procedures or outcomes, whereas only 4% of these evaluations focused upon the goals of intervention. The data for normative comparisons reflect a similar pattern to those for subjective evaluations. Postintervention assessments were slightly more frequent than preintervention assessments. In regard to preintervention normative comparisons, 86% concerned the goals of the experiment. Postintervention normative comparisons occurred only for the outcomes of intervention.

In general, the reporting of postintervention subjective assessments appears to be the primary type of social validity measure employed. Of the 198 instances in which subjective evaluations were employed, 155 were conducted after the experiment was completed. It appears from the data in Table 1 that subjective evaluations focusing upon the procedures used and the behavioral outcomes re-

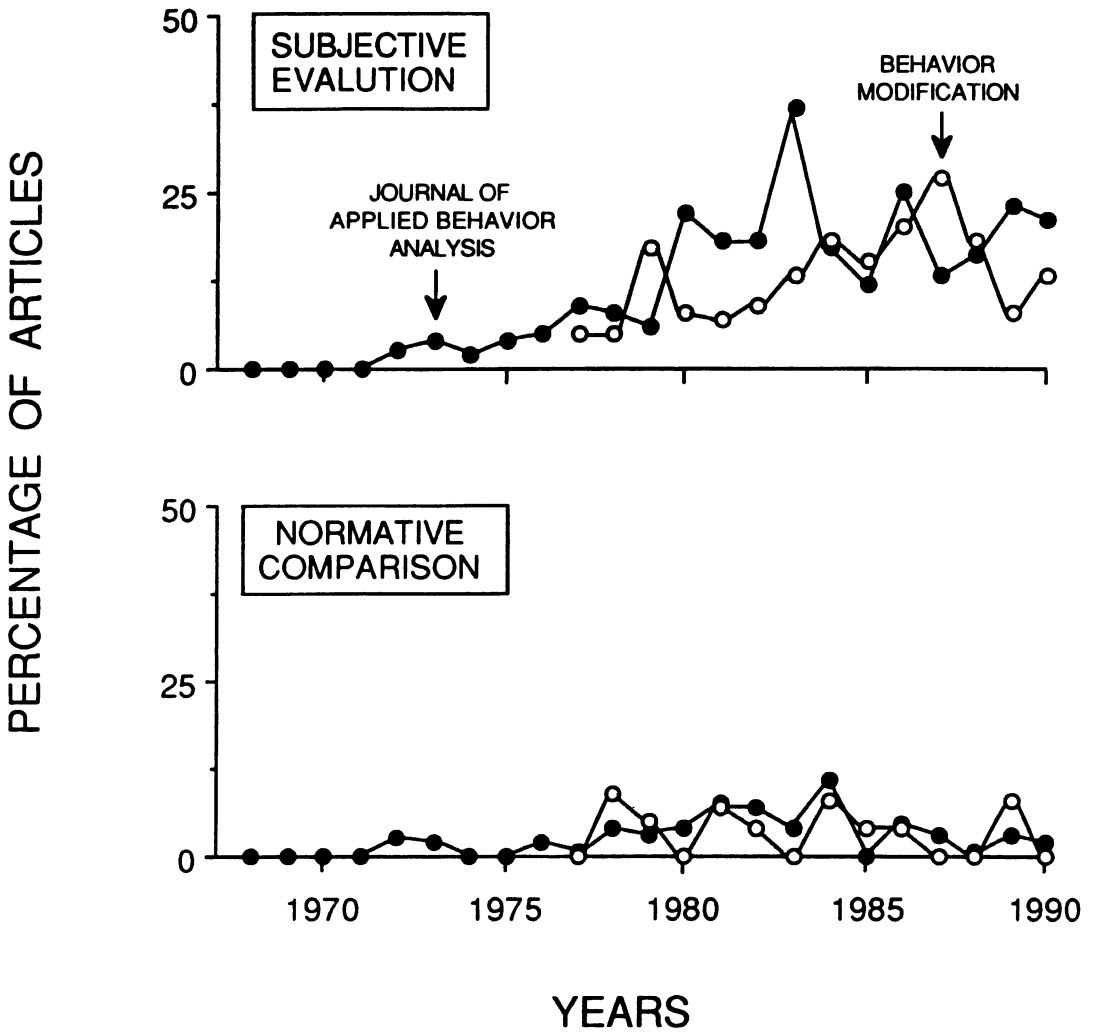

Figure 2. The use of subjective evaluation and normative comparison as social validity measures.

sulting from the intervention(s) were the most prevalent forms.

Use of Multiple Social Validation Strategies

Also of interest is the use of multiple measures to assess social validity; that is, the combined use of either multiple types of subjective evaluation, normative comparison, or both. Few articles reported the use of multiple social validation measures prior to publication of the Kazdin (1977) and Wolf (1978) articles. The use of multiple measures peaked in 1983 and has leveled off and remained constant since that time at approximately 10% of all articles published (primarily combinations of subjective evaluation) (data not

TABLE 1

A summary of the use of social validity strategies in *JABA* and *Behavior Modification*. The data are presented as the number of times a particular social validity strategy has appeared in a study

	Goals	Procedures	Outcomes
Subjective evaluation			
Preintervention	34	6	3
Postintervention	6	44	105
Normative comparison			
Preintervention	18	1	2
Postintervention	0	0	16

shown). This pattern parallels the overall use of social validity measures. The similarities indicate that the use of multiple measures has covaried with increases and decreases in the overall use of social validity.

DISCUSSION

Overall, the use of social validity measures appears to have been restricted to approximately one fifth of research articles published in *JABA* and *Behavior Modification,* relying primarily upon the subjective evaluation of outcomes after intervention. Although the use of social validity measures by applied behavior analysts stabilized during the 1980s at 20% of all data-based reports in the two journals, questions remain regarding how often these measures are used and the range of phenomena that are assessed.

Recently, the need to advance our understanding of social validity has been discussed (see the special issue of *JABA, 24*(2), 1991). Specific areas have been identified in which current social validity measures need to evolve. These areas include (a) determining what type(s) of articles should provide social validity indexes, (b) whether an increase in the use of social validity measures and/or development of novel approaches to social validity can be brought about by changes in editorial policy, and (c) the adequacy of social validity measures in assessing the effects of interventions in social contexts (Fawcett, 1991; Hawkins, 1991; Schwartz & Baer, 1991). Each of these issues raises questions not only about the utility of social validity measures but also about how we define and conceptualize this construct. This paper provides quantitative information from which each of these issues can be considered further.

In regard to the proportion of articles in *JABA* and *Behavior Modification* presenting social validity data, a stable pattern currently exists in which one out of every five data-based articles incorporates this type of information into the analysis. These figures suggest that the majority of applied articles providing original data do not seek out social validation. Should all experimental reports regarding questions of applied significance present social validity measures? This question raises the issue of discriminating between articles having as their primary goal the improvement (either through increasing and/or decreasing behavior) of an area of social concern versus articles primarily serving a research and development function. In other words, should articles explicitly attempting to remedy societal problems be contrasted with research analyses designed to study basic behavioral functions and processes in naturalistic settings? Both categories of research appear frequently. Reports from the former category should provide assessments of the social impact of interventions. Reports analyzing basic relations between behavior and naturalistic environments should be explicit in stating the goals of the experiment and need not, in the majority of cases, provide social validity assessments (e.g., Haring & Kennedy, 1990; Horner & Day, 1991; Mace, McCurdy, & Quigley, 1990; Wacker et al., 1990).

For applied investigations focusing explicitly on remedying problems of social concern, the question is raised regarding how to increase the proportion of articles presenting social validity data. One approach is to adopt an editorial policy specifying that all articles published must present social validity outcomes. Mandating 100% compliance with such a policy, however, may conflict with the distinction made previously: Not all applied research reports are necessarily conducted to solve social problems immediately. Thus, the presentation of social validity data as a prerequisite for publication consideration may be too stringent a criterion. A second approach to increasing the use of social validity measures may be the development of new approaches (e.g., Winett et al., 1991). By increasing the range of options available to researchers in regard to social validation, a corresponding increase in the use of these measures may occur. The information presented here clearly indicates a very restricted use of social validity measures in current practice. Developing social va-

lidity practices that better address the range of applied concerns may be one means of increasing its use.

Trends in the measurement of social validity may also be a reflection of the insensitivity of current measures to the "ecological" variables we are interested in quantifying (Baer, 1986). One way to view the sensitivity of a measure to changes in behavior is the measure's relation to the contingencies acting upon behavior. As stated by Kazdin (1978) in discussing the role of measurement sensitivity in applied analyses, "The overriding assumption of the operant approach is that behavior is a function of its consequences. . . . If consequences for a given activity are discontinued [or altered] . . . behavior would be expected to adjust to these contingencies" (p. 298). If we measure aspects of behavior closely related to the contingencies present in a social context, those measures should covary as the contingencies change or are altered. Perhaps the use of subjective perceptions regarding a specific aspect of an intervention has not provided researchers with a sensitive enough measure of the contingencies present in social contexts. Which box a teacher or parent checks on a 5-point Likert-type scale in reply to the question "Does the child show an interest in toys?" may so constrain informants' responses that the social importance of a particular interventions is not being captured. Similarly, comparing normative rates of one population in one particular context and time with another populations in another context and time may not provide researchers with the information needed. Baer, Wolf, and Risley (1987) observed that "social validity is sometimes assessed at present in very rudimentary ways that may too often find social validity where it does not actually operate" (p. 333).

The relation between subjective evaluations or normative comparisons and how behavior change affects a social ecology may be too remote to reflect many changes of experimental interest. In fact, the changes may not be readily apparent or necessarily identifiable to the individuals in a context—including parents, teachers, students with or without disabilities, behavior-change agents, and/or behavior analysts. It has long been recognized that what people perceive, how they subjectively experience those perceptions, and the actual events that occur sometimes do not correspond (e.g., Skinner, 1956). The tactic of developing measures that more sensitively reflect behavioral changes in social contexts may provide an important alternative direction for expanding our understanding of social validity.

CONCLUSION

Applied behavior analysis appears to be at an important point in regard to the use of social validity. Current uses of social validity, based upon the content analysis presented here, are based primarily on assessment of subjective perceptions of experimental outcomes. Recent attention given to the conceptualization and measurement of social validity, however, may change the trends presented in this paper. One approach may be to differentiate between applied analyses that have as a goal the improvement of behavior of social concern, and analyses that serve a basic research and development function but that do not focus directly on ameliorating problems of social concern. It has been suggested here that not all analyses published in journals such as *JABA* and *Behavior Modification* should provide data regarding social validity. Rather, researchers should work toward a viable means of differentiating among those reports in which measuring social validity is a necessary means of demonstrating the applied utility of the investigation and those reports in which social validity measures are not necessary. Perhaps by distinguishing between the necessary and unnecessary instances for the measurement of social validity, the foundation can be established for evaluating both *when* to use social validity measures and *what* measures are most appropriate. Increasing the diversity and frequency of social validation practices, along with a differentiation among various goals of applied behavioral research,

may provide an important step toward improving the effectiveness and durability of applied interventions.

REFERENCES

Baer, D. M. (1974). On the absence of Santa Claus in any known eco-system. *Journal of Applied Behavior Analysis, 7,* 167–171.

Baer, D. M. (1986). In application, frequency is not the only estimate of the probability of behavior units. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 117–136). Hillsdale, NJ: Lawrence Erlbaum.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 20,* 313–327.

Barahal, D. G., Bramer, L. M., & Shostrom, E. L. (1950). A client-centered approach to education-vocational counseling. *Journal of Consulting Psychology, 14,* 256–260.

Barlow, D. H. (1981). On the relation of clinical research to clinical practice: Current issues, new directions. *Journal of Consulting and Clinical Psychology, 49,* 147–159.

Bordin, E. S. (1955). The implications of client expectations for the counseling process. *Journal of Counseling Psychology, 2,* 17–21.

Bornstein, P. H., & Rychtarik, R. G. (1983). Consumer satisfaction in adult behavior therapy: Procedures, problems, and future perspectives. *Behavior Therapy, 14,* 191–208.

Brigham, T. A., Graubard, P. S., & Stans, A. (1972). Analysis of the effects of sequential reinforcement contingencies on aspects of composition. *Journal of Applied Behavior Analysis, 5,* 421–434.

Briscoe, R. V., Hoffman, D. B., & Bailey, J. S. (1975). Behavioral community psychology: Training a community board to problem solve. *Journal of Applied Behavior Analysis, 8,* 157–165.

Brookes, P. H., & Baumeister, A. A. (1977). A plea for consideration of ecological validity in the experimental psychology of mental retardation. *American Journal of Mental Deficiency, 81,* 407–416.

Charlop, M. H., & Milstein, J. P. (1989). Teaching autistic children conversational speech using video modeling. *Journal of Applied Behavior Analysis, 22,* 275–285.

Clark, H. B., Greene, B. F., Macrea, J. W., McNees, M. P., Davis, J. L., & Risley, T. R. (1977). A parent training package for family shopping trips: Development and evaluation. *Journal of Applied Behavior Analysis, 10,* 605–624.

Elliot, S. N., Witt, J. C., Galvin, G. E., & Moe, G. L. (1986). Children's involvement in intervention selection: Acceptability of interventions for misbehaving peers. *Professional Psychology: Research and Practice, 17,* 235–241.

Fawcett, S. B. (1991). Social validity: A note on methodology. *Journal of Applied Behavior Analysis, 24,* 235–240.

Fawcett, S. B., & Miller, K. L. (1975). Training public speaking behavior: An experimental analysis and social validation. *Journal of Applied Behavior Analysis, 8,* 125–141.

Friman, P. C., Finney, J. W., Rapoff, M. A., & Christophersen, E. R. (1985). Improving pediatric appointment keeping with reminders and reduced response requirement. *Journal of Applied Behavior Analysis, 18,* 315–321.

Gaylord-Ross, R. J. (1979). Mental retardation research, ecological validity, and the delivery of longitudinal educational programs. *Journal of Special Education, 13,* 69–80.

Goode, W. F., & Fowler, I. (1949). Incentive factor in a low morale plant. *American Sociological Review, 14,* 79–93.

Grant, C. W. (1954). How students perceive the counselor's role. *Personnel and Guidance Journal, 38,* 386–394.

Green, R. B., Hardison, W. L., & Greene, B. F. (1984). Turning the table on advice programs for parents: Using placemats to enhance family interaction at restaurants. *Journal of Applied Behavior Analysis, 17,* 497–507.

Gullone, E., & King, N. J. (1989). Acceptability of behavioral interventions: Child and caregiver perceptions. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 24, pp. 132–151). Newbury Park, CA: Sage.

Haring, T. G., & Kennedy, C. H. (1990). Contextual control of problem behavior in students with severe disabilities. *Journal of Applied Behavior Analysis, 23,* 235–243.

Hawkins, R. P. (1975). Who decided that was the problem? Two stages of responsibility for applied behavior analysts. In W. S. Wood (Ed.), *Issues in evaluating behavior modification* (pp. 195–214). Champaign, IL: Research Press.

Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis, 24,* 205–213.

Herzberg, F., Mausner, B., & Syndern, B. B. (1959). *The motivation to work.* New York: John Wiley & Sons.

Horner, R. H., & Day, H. M. (1991). The effects of response efficiency on functionally equivalent competing behaviors. *Journal of Applied Behavior Analysis, 24,* 719–732.

Kazdin, A. E. (1977). Assessing the clinical or applied significance of behavior change through social validation. *Behavior Modification, 1,* 427–452.

Kazdin, A. E. (1978). *History of behavior modification: Experimental foundations of contemporary research.* Baltimore, MD: University Park Press.

Kazdin, A. E. (1980). Acceptability of alternative treatments for deviant child behavior. *Journal of Applied Behavior Analysis,* 13, 259–273.

Kazdin, A. E., & Wilson, G. T. (1978). Criteria for evaluating psychotherapy. *Archives of General Psychiatry, 35,* 407–416.

Koos, E. L. (1955). "Metropolis"—What city people think of their medical services. *American Journal of Public Health, 45,* 1551–1557.

Lebow, J. L. (1982). Consumer satisfaction with

mental health treatment. *Psychological Bulletin, 91,* 244–259.

Lovitt, T. C. (1978). New applications and new techniques in behavior modification. *Journal of Special Education, 12,* 89–93.

Mace, F. C., McCurdy, B., & Quigley, E. A. (1990). A collateral effect of reward predicted by matching theory. *Journal of Applied Behavior Analysis, 23,* 197–206.

Makeover, H. B. (1950). The quality of medical care. *American Journal of Public Health, 41,* 824–832.

Maloney, K. B., & Hopkins, B. L. (1973). The modification of sentence structure and its relationship to subjective judgments of creativity in writing. *Journal of Applied Behavior Analysis, 6,* 425–441.

McMahon, R. J. (1983). Consumer satisfaction in behavioral treatment of children: Types, issues, and recommendations. *Behavior Therapy, 14,* 209–225.

Nutter, D., & Reid, D. H. (1978). Teaching retarded women a clothing selection skill using community norms. *Journal of Applied Behavior Analysis, 11,* 457–487.

O'Brien, F., & Azrin, N. H. (1972). Developing proper mealtime behaviors of the institutionalized retarded. *Journal of Applied Behavior Analysis, 5,* 389–396.

Phillips, E. L., Wolf, M. M., & Fixsen, D. L. (1973). Achievement Place: Development of the elected manager system. *Journal of Applied Behavior Analysis, 6,* 541–554.

Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker.* Cambridge, MA: Harvard University Press.

Rogers, C. R. (1942). *Counseling and psychotherapy: New concepts in practice.* New York: Houghton Mifflin Company.

Rogers-Warren, A., & Warren, S. F. (1977). *Ecological perspectives in applied behavior analysis.* Baltimore, MD: University Park Press.

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis, 24,* 189–103.

Scott, M. (1980). Ecological theory and methods for research in special education. *Journal of Special Education, 14,* 279–294.

Singh, N. N., & Katz, R. C. (1985). On the modification of acceptability ratings for alternative child treatments. *Behavior Modification, 9,* 375–386.

Skinner, B. F. (1956). A case history in scientific method. *American Psychologist, 11,* 221–233.

Stevenson, H. C., & Fantuzzo, J. W. (1986). The generality and social validity of a competency-based self-control training intervention for underachieving students. *Journal of Applied Behavior Analysis, 19,* 269–275.

Van Houten, R. (1979). Social validation: The evolution of standards of competency for target behaviors. *Journal of Applied Behavior Analysis, 12,* 581–591.

Van Houten, R., Morrison, E., Jarvis, R., & MacDonald, M. (1974). The effects of explicit timing and feedback on compositional response rate in elementary school children. *Journal of Applied Behavior Analysis, 7,* 547–554.

Wacker, D. P., Steege, M. W., Northup, J., Sasso, G., Berg, W., Reimers, T., Cooper, L., Cigrand, K., & Donn, L. (1990). A component analysis of functional communication training across three topographies of severe behavior problems. *Journal of Applied Behavior Analysis, 23,* 417–430.

Wienerman, E. R. (1964). Patients' perceptions of group medical care: A review and analysis of studies on choice and utilization of prepaid group practice plans. *Group Medical Care, 54,* 880–889.

Willems, E. P. (1974). Behavioral technology and behavioral ecology. *Journal of Applied Behavior Analysis, 7,* 151–165.

Winett, R. A., Moore, J. F., & Anderson, E. S. (1991). Extending the concept of social validity: Behavior analysis for disease prevention and health promotion. *Journal of Applied Behavior Analysis, 24,* 215–230.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11,* 203–214.

Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strengths, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49,* 156–167.